2. The instrument basic characteristics, graph and the least squares method

2.1 Characteristics of measurement system

It is necessary to know representation of performance of measurement system. But it does not mean that uniform definition applies on different measurement devices. Use modern definition may have some different with conventional definition.

2.1.1 Static characteristics

(1) Sensitivity

Sensitivity can show the degree of change of measured quantity. Generally, representation of static variation at input side and output side.

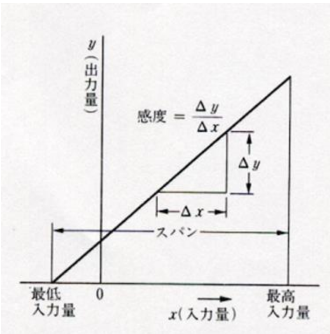

Fig.2.1 Sensitivity

Sensitivity is defined as a ratio between input change amount and output change amount, and given by slope. For example, when input is 1[mm], and pointer move 1[mv], so 1 [mv/mm], also, input of digital voltmeter is 1 [μV], and display show 1 [digit], which means 1 [digit/μV]. But by using magnifying mechanism or amplifier, input and output still same dimension. In this case, a ratio of the input and output is different from sensitivity, called as enlargement ratio.

(2) Resolution

If input is not zero and increase slowly, when the increase is become some value, and output will begin to change. In this case, the minimum amount of increase in the input that can be detected is called resolution. It also called sensitivity limit. The device that using for collect substance information, call that with detection limit. Numeric display use percent (%) against absolute value or span.

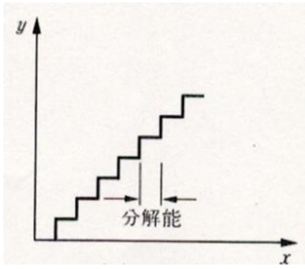

Fig.2.2 Resolution

(3) Measuring range, Span, Dynamic range

Measuring range is a range that measurable amount of measurement device. For example, daily use wall hanger hygrometer, temperature scale is $-20\sim40$[℃], relative humidity scale is $10\sim90\%$. Difference of minimum value and maximum value is called span (Fig.2.1). In the example, span is 60[℃] and 80[%].

To measurement device, less time to measure object correctly is desired. Measurement device respond dynamically, and dynamic range means ratio of maximum input and minimum sensing input (dynamic resolution), express by decibel [dB]. Maximum input is x_{max}, minimum sensing input is δ_x, dynamic range R_d can be expressed by,

$$R_d = 20log\frac{x_{max}}{\delta_x}[dB]$$

(4) Linearity

It is convenient in math that relation of input and output of measurement device is linearity. Relation of input and output is represented by non-linear data, defined by linearity. Usually, Linearity is expressed by ratio against span of maximum deviation of input. Disadvantage of non-linearity is, the bigger the non-linearity, the more the calibration points for the standard. In measurement device, there are built-in circuits or process for linearity. This correction mechanism is called linearizer.

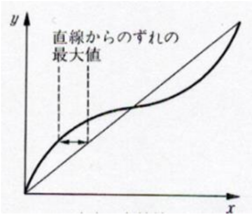

Fig.2.4 Linearity

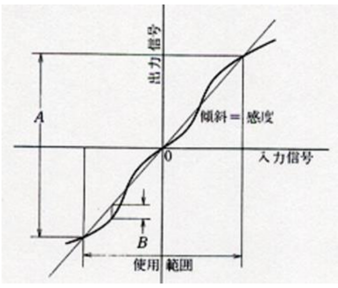

Fig.2.5 Relation of input and output

(5) Hysteretic Error, Accuracy

Hysteretic Error is difference of same input caused by history of measurement and it can be shown in Fig.2.6. Accuracy is the error limit of the value displayed by the measurement device under certain specified conditions, and shows performance of instrument.

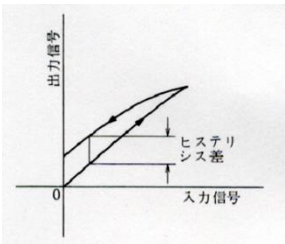

Fig.2.6 Hysteretic Error

2.1.2 Dynamic characteristics

(1) Step response

As shown in Fig.2.7, in the case of input is step signal, when output error comes to less than a value, the required time is called response time. Response time is, characteristics of measurement device that track intense change of input for responds it on output correctly. System that increase rapidly, the take time that 10~90 [%] of final output is called rise time.

On the other hand, in Fig.2.8, for a system that change of output is like an exponential function, the cost time that 63.2% of final output is called time constant.

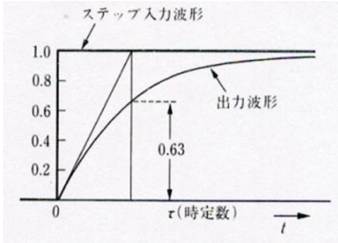

Fig.2.7 Response time

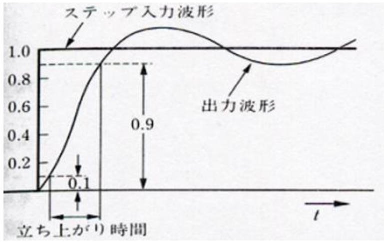

Fig.2.8 Time Constant

(2) Frequency response

In steady state, amplitude of input signal is A, amplitude of output is B was supposed, the examples can be shown in Fig.2.9 and 2.10. Frequency response is a measuring equipment output characteristic that the value B/A and phase difference φ of the depending on angular frequency ($\omega=2\pi f$), where f is frequency.

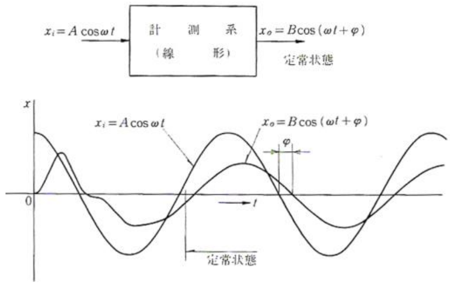

Fig.2.9 response of sine wave

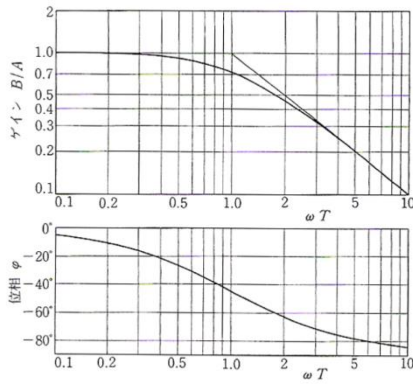

Fig.2.10 variation of B/A and φ

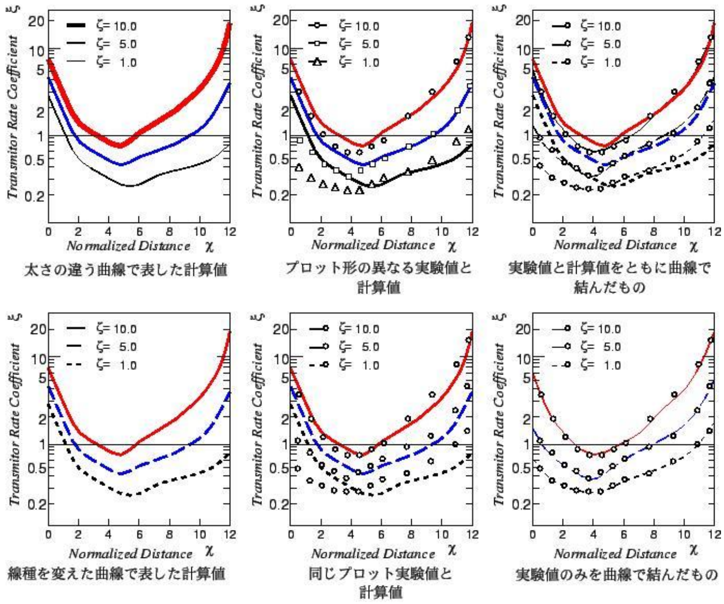

Fig.2.11 Example of graph

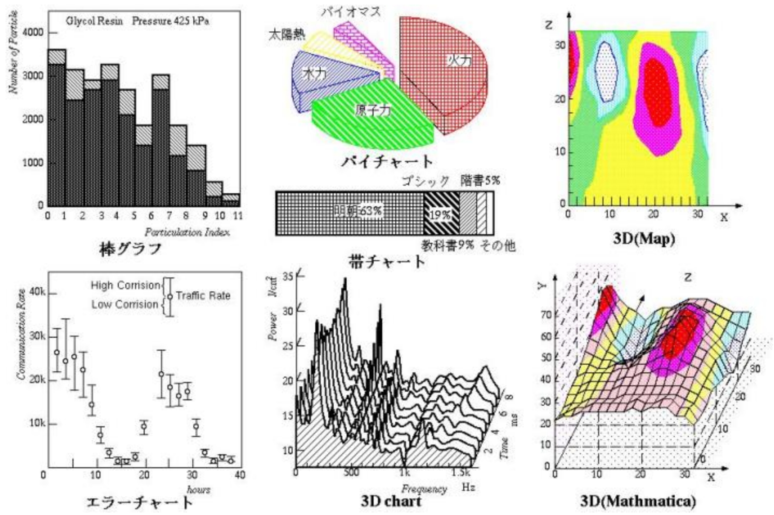

Fig.2.12 Types of graph

2.2 Graph and Least squares method

2.2.1 Graph

A graph is a tool for expressing data, human beings to analyze data, and also a tool for explaining to others. The created graph is strongly likely to be reused if it expresses important properties. For this reason, it must extract features so that intuitive visibility is given to the content to be explained. Even with exactly the same data, expressive power differs if you change the line or plot to use.

Fig.2.11 shows the example of graphs, where six graphs are shown in the order from left (a) to (f) above. In this figures, (a) and (d) emphasize calculated values of ζ with thickness, line type and color, (b) emphasizes the experiment value plot from the calculated value, (e) emphasizes calculated values. Figure (f) is a method to show changes to experimental values in an easy-to-understand manner. (c) is an example of a bad graph. In this way, even if you write a graph with the same data, the consciousness (sense) of the person who writes appears and you can tell the intention to the viewer.

Fig.2.12 shows the types of graph. As is shown in figure, types of graph are infinite when including business purposes such as bar graphs, pie charts, band graphs, etc. Of course, they are also used in engineering papers and presentations.

Generally, for 2D and 3D graphs, independent variables are arranged on the horizontal axis and dependent variables are arranged on the vertical axis. In other words, the horizontal axis represents the manipulated variable and the vertical axis represents the output quantity, and the meanings of the measured quantities and calculation results expressed in the graph are easy to understand. There are two ways to express this expression.

(1) Display original data directly

A graph created to intuitively grasp the physical quantity without applying hypotheses or theory, especially to experiments and measurement objects.

(2) Display processed data

A graph based on theoretical and hypothetical models for the characteristics of the object to be measured and displaying data on the extent corresponding to that model.

Writing a graph is nothing short of displaying ideas on data and data processing, but also represents the validity of the processing method.

2.2.2 Least squares method

Generally, linear relation of input and output is wanted. Even the relation of input and output is not linear, it is still necessary that the relation can be expressed with a function. This section discuss situation that the relation of input and output is linear, and use a method called least squares method. Even, the relation is not a linear; the relation may be changed to linear by calculating scale with logarithm. So, this method is extremely useful.

The least squares method is a method of calculating the most accurate and reliable value as a function value by minimizing the sum of squares of residuals, when there are a large number of measured values having the same degree of accuracy. If there are x, y can be written in

$$y = ax + b,$$

and there are N-dimension measured data, $(x_1, y_1), (x_2, y_2), \cdots, (x_3, y_3)$. The problem is to solve the best liner to fit the measured data. To solve the problem, there is no way but to calculate a, b. The finest linear is the a, b that make the sum of the squares of the residuals of y_i and $ax_i + b$ minimum.

$$\sum_{i=1}^{n} [y_i - (ax_i + b)]^2$$

Find the a, b which make S minimum. Which means find a, b at differential is 0.

$$\frac{dS}{da} = 2\sum_{i=1}^{n} \{[y_i - (ax_i + b)](-x_i)\}$$

$$= 2[-\sum x_i y_i + a\sum x_i^2 + b\sum x_i] = 0.$$

Which also can be written in

$$a\sum x_i^2 + b\sum x_i = \sum x_i y_i$$

And in same way $\frac{dS}{da} = 0$

$$a\sum x_i + nb = \sum y_i$$

The a, b can be written in

$$a = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{n\sum x_i^2 - (\sum x_i)^2}$$

$$b = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i y_i)(\sum x_i)}{n\sum x_i^2 - (\sum x_i)^2}$$

Example:

a) Use least squares method to solve this problem:

x_i	0.75	1.5	2.8	4.0	5.0
y_i	1.0	1.6	2.0	2.8	3.2

Solution:

x_i	y_i	$x_i y_i$	x_i^2
0.75	1.0	0.75	0.563
1.5	1.6	2.40	2.250
2.8	2.0	5.60	7.840
4.0	2.8	11.20	16.000
5.0	3.2	16.00	25.000
$\sum x_i$ = 14.05	$\sum y_i$ = 10.6	$\sum x_i y_i$ = 35.95	$\sum x_i^2$ = 51.563

$$a = \frac{5 \times 35.95 - 14.05 \times 10.6}{5 \times 51.653 - 14.05^2} = 0.506$$

$$b = \frac{51.653 \times 10.6 - 35.95 \times 14.05}{5 \times 51.653 - 14.05^2} = 0.697$$

So, the linear is $y = 0.506x + 0.697$, and picture can be shown in Fig.2.11.

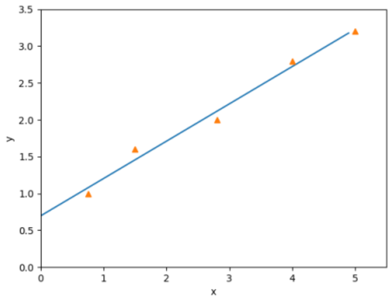

Fig.2.11 Result of example a)

b) If $y = ke^{ax}$, and x, y is been given. Use least squares method to solve k, a.

x	0	0.5	1.0	1.5	2.0
y	1.350	1.734	2.226	2.858	3.675

Solution:

Taking the logarithm of both sides

$$log_e y = ax + log_e k$$

Calculate $log_e y$ rather than y, the same way like example a).

$$a = 0.501$$

$$log_e k = 0.300$$

$$k = 1.349$$

```python
import matplotlib.pyplot as plt
import numpy as np

x = np.array([0.75, 1.5, 2.8, 4.0, 5.0])
y = np.array([1.0, 1.6, 2.0, 2.8, 3.2])

if x.size != y.size or x.ndim != y.ndim:
    print("warning")

else:
    xy = x * y
    x_square = x ** 2

    x_sum = np.sum(x)
    y_sum = np.sum(y)
    xy_sum = np.sum(xy)
    x_square_sum = np.sum(x_square)
    n = x.size

    a = (n*xy_sum - x_sum*y_sum) / (n*x_square_sum - x_sum**2)
    b = (x_square_sum*y_sum - xy_sum*x_sum) / (n*x_square_sum - x_sum**2)

    def lsm(x):
        return a*x + b

    plt.xlabel("x")
    plt.ylabel("y")

    xl = np.arange(np.amin(x), np.amax(x), 0.1)
    yl = lsm(xl)

    plt.plot(xl, yl)
    plt.plot(x, y, '^')
    plt.show()
```